Part II

Applications of Cash on Delivery Aid

Applying COD Aid to primary education: a contract between partners

Cash on Delivery Aid (COD Aid) could be applied in many sectors and in a variety of financial relationships. But is it practical? In this chapter we explain how foreign aid agencies could use COD Aid to collaborate with developing countries in reaching universal primary schooling. Applying COD Aid to this specific example illustrates the kinds of problems to be addressed and demonstrates that most of them are manageable. The challenges seem no greater than those that arise with traditional aid modalities. (In chapter 4 we turn to practical issues of funding and implementation.)

Essential elements of a COD Aid agreement for primary education

As chapter 2 discussed, COD Aid has five key features: payment for outcomes, hands-off funders and responsible recipients, independent verification, transparency through public dissemination, and complementarity with existing aid programs. In practical terms, this requires that a COD Aid contract include four essential elements:

- A shared and clearly defined goal.
- A unit for measuring progress.
- Payment per unit of progress.
- A system for measuring and verifying progress.

The first essential element is a goal that is shared by both funder and recipient and is clearly defined. If the goal is not genuinely shared, or is ancillary to other hidden objectives (such as tied aid, geopolitical allegiance of the funder, or other domestic priorities of the recipient), there is no point in embarking on a COD Aid agreement. Especially with full recipient discretion and responsibility, little progress will be made unless the goal merits real commitment from both parties. The implementation of the agreement

A COD Aid contract includes a system for measuring and verifying progress

also requires that the goal be sufficiently specified to permit accurate measurement and verification.

The second element of a COD Aid contract is a unit for measuring progress toward the shared goal. The unit of measurement should be relevant, precise, and capable of capturing continuous or incremental progress.

A definite statement of the amount to be paid for each unit of progress is the third essential element. That amount is not pegged to the cost of progress—it should be in proportion to the amount needed to attract the attention, energy, and commitment of high-level policy actors.

The final element of any COD Aid contract is a system for measuring and verifying progress. That system includes the means for a recipient to collect data on progress, the data the recipient collects, and the provision for third-party verification of data.

This chapter describes a hypothetical COD Aid agreement in the education sector. In this example, the four essential elements are embodied as follows:

- The shared goal is to ensure that every child completes primary education of good quality.
- The unit of progress is an "assessed completer," a student who is enrolled in the last year of primary school and who takes an approved standardized test.
- The funder agrees to pay $20 for each student who takes a standardized test in the last year of primary school up to the total enrollment in the base year and $200 for each assessed completer in excess of that number.
- The recipient commits to disseminating its information on student enrollments, assessed completers and test scores (at some agreed level of disaggregation). The funder commits to contracting a third party to verify the accuracy of the recipient's reports.

Box 3.1 provides a summary of the proposed contract. The rest of the chapter discusses the rationale behind the choices in preparing this contract. The appendix contains term sheets that could be used to negotiate such a contract.

Shared goal: universal primary completion

A good example of a shared goal in a COD Aid contract is to ensure that every child completes primary education of good quality. Universal primary education is an appropriate goal for such an agreement because it has already been endorsed by many countries, both in international agreements and in domestic policies. In September 2000, 189 countries signed the UN Millennium Declaration, which calls for universalizing primary school. Universal primary education is one of the eight Millennium Development Goals subsequently agreed.[1] Most countries have constitutional provisions or at least national legislation guaranteeing their population's access to primary schooling.

BOX 3.1

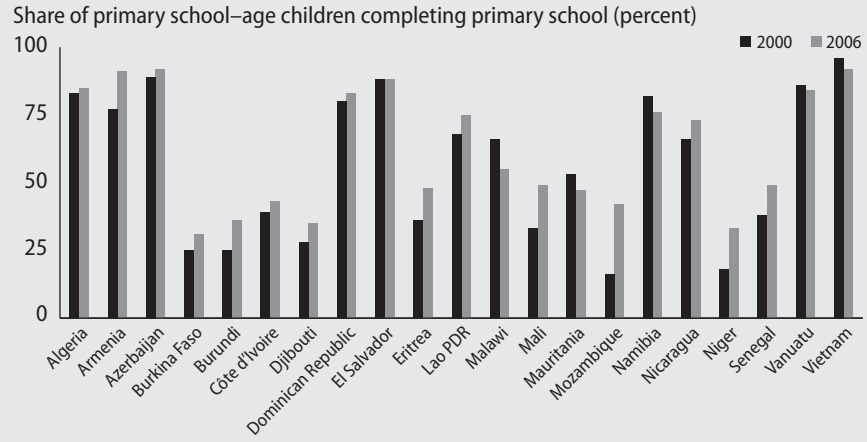
Outline for a COD Aid contract to pay for assessed completers

1. The funder commits to pay the recipient government $20 per assessed completer up to the total number of children enrolled in the last year of primary school in the base year and $200 per assessed completer beyond that number.
2. The contract is for five years, extendable in five-year increments.
3. Only students enrolled in the last year of primary schooling are eligible to be counted as assessed completers for the COD Aid agreement.
4. The funder will pay only once for each student, regardless of how many times he or she may repeat or retake the test.
5. The funder will contract an agent from a preapproved list of organizations to verify the government's report on the number of assessed completers and their scores.
6. The agent will retest a randomly selected sample of schools and disseminate the findings publicly.
7. If the retest findings show the official reports are accurate within a 5 percent margin, the COD payment will be calculated as described in item 1. Otherwise, the COD payment will be reduced to reflect the agent's estimate along with a penalty.
8. The participating funder commits to ensuring that the COD payments are treated as additional to other assistance to the country.
9. At the inception of the contract, the funder will place a guarantee in escrow equivalent to the amount that would be disbursed over the subsequent two years if the country were able to test 90 percent of the base year enrollment in the last year of primary school. The funder will replenish the escrow account annually so that, in any given year, the account holds enough funds for the subsequent two years based on mutually agreed projections.
10. Funders will designate an account to which countries may charge up to 90 percent of the direct costs of developing a robust information system on student completion and of developing and rolling out the standardized competency test.
11. The standardized competency test will be designed to allow accurate tracking of learning outcomes from year to year in order to assess whether schooling quality is improving and to assist management and education policy decisions.
12. The recipient commits to publicly report student completion figures and average test scores, with levels of disaggregation appropriate to the accuracy and reliability of the test, to be agreed as a part of the contract.
13. The recipient commits to allow and facilitate research into education policy, the development of institutions, and the effects of the contract, and to make complete individual-level education data and public finance data available to researchers for this purpose, with appropriate measures taken to protect the privacy of all individuals.
14. The funders and recipient agree that, in the event of a dispute, they will follow procedures to select and abide by the decisions of an international arbitration panel.

Despite the shared commitment to reaching the Millennium Development Goal of universal primary completion by 2015, many countries are unlikely to reach it (figure 3.1). COD Aid is a way to reinvigorate efforts and accelerate progress in these

FIGURE 3.1

Many countries are unlikely to achieve the Millennium Development Goal of universal primary completion by 2015



Source: World Bank 2008.

countries; one reviewer even described the COD Aid proposal as "Millennium Development Goals with teeth."

Measure of progress: assessed completers

The next element is to identify an indicator that would measure progress toward the shared goal. In this example, we propose that progress be measured by the number of assessed completers—that is, the number of students enrolled in the last year of primary school who complete a standardized test.[2] This measure has several advantages:

- It is easily defined.
- It is closely related to the goal of universal primary education.
- It draws attention to the content and quality of schooling.
- It is relatively easy to audit.
- It encourages the development of management information and testing systems.

Number of assessed completers is easily defined

Why count the number of students taking a test to measure increases in completion when school records could be used instead? First, the definition of completing primary school is problematic. Many countries lack a clear definition of what is required to complete the cycle in terms of minimum attendance, learning standards, or performance on any formal system of assessment for progressing through the school system and graduating. As a result, most international data on the number of children who complete primary school are little more than the difference between the number of

students who enroll in the last year of primary school minus the number who are held back to repeat that year.[3] Countries that lack a clear definition of primary school completion would need to negotiate such a standard with the funder, and verification would require a detailed audit to assess whether the standard has been met. Such an undertaking would be expensive, requiring an audit of attendance records among other things. By contrast, the number of students taking a test is clearly defined and fairly simple to measure.

Number of assessed completers is related to the goal of universal primary education

An assessed completer is not exactly equivalent to ensuring that every child completes a good quality primary education, for a variety of reasons. Students may be absent for a large number of school days or be poorly taught. Even so, the test scores will give some indication of the share and number of students who have mastered the educational content that is expected of students who complete primary school.

Number of assessed completers draws attention to the content and quality of schooling

If no test were administered, recipients might be tempted to expand enrollment at the expense of school quality. By paying for the number of last-year students who take a test, the proposed COD Aid contract ensures that the quality of schooling is not ignored. The requirement that these test scores be published—disaggregated to a level that is appropriate and feasible—draws attention to the content of schooling and the performance of schools. This is likely to facilitate public debate about the quality and equity of the education system. Public dissemination also encourages the government to respond more to concerns of civil society than to concerns of funders. Furthermore, the education community will be able to use the published test scores in interpreting results and improving the assessment system. Although our proposed indicator does not make payment contingent on achieving any individual or aggregate test scores, public dissemination of the results is required and should help make government and school systems more accountable for progress on quality as well as quantity of schooling. Over time, this may be one of the most valuable aspects of the agreement.

Public dissemination of results should make government and school systems more accountable for progress

Number of assessed completers is relatively easy to audit

Verifying the number of assessed completers is much easier than verifying the number of students who complete primary school. Auditors need only check the number of students who took a test on a particular day, rather than verify that students have completed some minimum attendance over a long period. The tasks, scale, and duration of the audit can be precisely defined, and the opportunities for manipulating and cheating

A COD Aid contract encourages countries to improve their administrative information

are reasonably limited. In addition, the audit can take place a week or two after the initial test, providing timely feedback on the accuracy of the indicator that will trigger payments (see further details below under reporting and verification).

Number of assessed completers encourages the development of management information and testing systems

A COD Aid contract that pays for assessed completers encourages countries to administer externally validated tests that give useful feedback to national policymakers, state governments, school districts, schools, teachers, students, and parents. It encourages countries to improve their administrative information so they can better track who is and is not in school, fostering the development of domestic capacity to assess students and analyze student performance. The benefits of testing and publishing data also mean that even if the COD Aid agreement results in few additional students completing school, the aid has still accomplished something of substantial value: improving the evidence base for effective policymaking.

A standardized test of learning provides policymakers and parents with useful and comparable information.[4] The COD Aid contract to pay for assessed completers could be associated with any particular test that is mutually agreeable to the funder and the recipient. But to reflect the shared goal of reaching universal primary school completion, the test should relate to the learning expected at that level. The most appropriate and feasible test for such a purpose would probably be a standardized test of competencies or skills (such as basic reading, writing, and mathematics) commonly acquired on completing primary school. The test should also be able to discern changes in learning from cohort to cohort over time.

To be useful for education policy, such standards-based assessments do not need to be comparable across countries. But they do need to be stable over time in a number of ways. They should test the same age cohorts and measure the same content or competencies. Measurement instruments should be at the same levels of difficulty and reliability. Empirical equating—ensuring that the tests are measuring the same constructs, have the same reliability, and are population invariant—should be conducted to confirm the equivalence of the test content and results over time.[5]

Tests based on representative samples must use the same sampling procedure from year to year. Those based on a full student census need to collect enough information to control for the changing composition of students from year to year. And the country must have the capacity to scientifically design, administer, score, and analyze the test or must be willing to outsource or purchase services that it cannot itself provide.

Few developing countries fulfill all these criteria with current national assessments. At most, 20 low- and middle-income countries conduct annual national learning

assessments. The analysis needed to judge the reliability of the tests is not generally completed, and few tests have been calibrated through empirical equating. Many low- and middle-income countries are making substantial progress toward routine application and use of national learning assessments, but they might find it challenging to design and implement such a test in the short run. They could use or adapt an existing international or regional test that meets the same standards while beginning the process of developing their own national test to meet these standards, or could choose to rely on existing international and regional tests for the long run.[6]

Deciding exactly which dimensions to measure is a valuable part of national policy debates

To avoid intrusiveness with assistance in test development, the recipient should choose whether and which firm or agency to contract with for technical support. The firm or agency could be selected from a number of organizations with appropriate expertise, including testing firms, international agencies such as the United Nations Educational, Scientific, and Cultural Organization, and organizations that have implemented international and regional learning assessments in recent years.[7] This is the one activity for which we recommend that the funder provide financial resources upfront, but the choice of whether and with whom to seek assistance remains the recipient's responsibility (box 3.2).

Testing must measure what society wants children to learn

Testing students serves many purposes.[8] Sometimes it is meant to select students for placement in subsequent levels of instruction or jobs; sometimes it is used to assess student performance as feedback for the student, parents, and teacher. At other times it is aimed at assessing the quality of the schooling system itself. All testing can be controversial. Education is a complex and multifaceted process, and no test can possibly measure all its various dimensions. Even so, properly executed tests can measure some aspects of education.

The process of deciding exactly which dimensions to measure is itself a valuable part of national policy debates. Introducing national tests, even when rewards and penalties are not associated with results, is likely to encourage teachers and schools to teach material and skills that are tested—that is, to teach to the test. In itself this is not a bad thing so long as the test is valid and reliable and measures the key outcomes that a society expects from its children's education. This practice becomes a serious problem only when the test becomes the sole focus of education to the exclusion of other less easily measured aspects of schooling.

But be careful how you use tests due to weak validity, high stakes, and cheating

Once a test is required, it is tempting to condition the payment on student test performance. But our consultations with experts and review of previous experiences

BOX 3.2

Paying upfront costs for test development and implementation

The COD Aid contract is designed to channel resources toward governments as they make progress on outcomes on which they and funders have already agreed to focus effort. This is why the COD Aid contract does not include advance funding for investments to expand and improve education. Funds from both foreign aid and domestic revenues should already be available for improving and expanding education; the additional COD Aid funds would flow as the government implements these programs and improves them. But the COD Aid contract does require one additional task for funds to flow: accurate reporting on outcomes.

Developing and implementing a national competency test, or upgrading an existing test to meet the standards described here, require upfront expenditure, as do assigning unique identifiers to students and upgrading existing education management information systems to enable accurate reporting of outcomes. The amount required will depend on the country and the systems already in place, but could range from a few hundred thousand dollars to a few million dollars.

Funders routinely provide aid to develop such systems because of the benefits for policymaking. Much of this aid is delivered in traditional input-driven ways, including packages of technical assistance to build local capacity to collect and manage the information. Such aid is often tied to providers of services and expertise from funding countries or at least to particular approaches. For reasons discussed earlier, much of this aid has failed. For example, much funding and expertise are invested without successfully equating a test or getting statistics to agree. But if efforts to develop stronger information systems were linked to the prospect of major flexible funding that would be available once a country qualified for cash payments, this might muster the focus and political will that could make the difference. Different parts of the bureaucracy might be pressured to get together and find out why their statistics disagree or wait to release test scores until they have been properly standardized.

We propose that funders incorporate a mechanism to help cover the direct upfront costs required to improve administrative reporting and information management enough to implement the COD Aid contract. Funders would agree to cover a certain percentage of eligible costs (say, 90 percent) with a predetermined ceiling and without requiring preapproval. The recipient could contract directly with providers of such services as test development and then send the bill to the funder.

have convinced us that this temptation has to be resisted. As discussed below, questions about test score validity and reliability will take time to resolve, and conditioning payments on scores that later prove to have been faulty could undermine the credibility of the entire project of testing. Meanwhile, staking large financial rewards on test performance encourages manipulation of results—by the recipient, schools, or even pupils. Such cheating increases the cost of auditing, reduces the utility of test scores for policymaking, and undermines efforts to introduce worthwhile student assessments. For these reasons, we propose that the payment be based on the number of children who take the exam and not on their test scores.

How many assessed completers are additional?

Which assessed completers would be paid for: all assessed completers or only those who would not have completed primary schooling without the COD Aid agreement? The ideal measure would be the difference between the actual number of assessed completers and a baseline calculated as a projection of how many students would have completed primary school without the COD Aid agreement. But any effort to project such a baseline would require extensive assumptions and the complications this entails would undermine the transparency of the agreement.

After reviewing this dilemma, we came to the conclusion that the baseline should be as simple as possible to facilitate transparency and reduce the risks of projecting a counterfactual future. Our proposal, therefore, is to calculate the number of additional assessed completers as the difference between the actual number of assessed completers and a baseline comprising the number of students who completed primary school in the first year of the program (base-year enrollment). If the agreement is extended beyond five years, as planned, then the baseline would be adjusted to the number of assessed completers five years earlier (the baseline would become completion with a five-year lag).

Adjusting the baseline is important for at least two reasons. First, it avoids a sharp drop in payments when universal completion is reached. Second, it strengthens the incentive to make additional effort to reach universal completion since every five years the payments associated with earlier achievements are removed.

The size of the payment

The COD Aid agreement has to establish how much the country will receive for each unit of progress. One approach is to calculate the cost of educating each student and use that as the basis for setting the payment. But the cost of educating students varies substantially across countries, across regions within countries, and across socioeconomic groups. Using the average cost could be seen as harmful to subpopulations or regions where the costs of schooling are higher than average. If, however, the fixed payment varied across countries, regions, or socioeconomic groups, it would seem to imply that the value placed on educating a child in one place is higher than for educating a child in another place. Above all, the idea of linking the payment to costs deflects attention from the real purpose of the payment: to be an incentive for the recipient to make progress toward what has already been agreed is a priority. Ultimately, countries are themselves committed to educating all their children and expect to do so in a financially sustainable fashion in the future. Thus, the COD Aid payment is not really aimed at covering the cost of schooling. It is aimed at relaxing constraints that hold back progress.

The COD Aid payment is aimed at relaxing constraints that hold back progress

$200 per assessed completer would get the attention of policymakers and engage them in improving efficiency

A payment of $200 per student could be sufficient incentive

After consulting with a wide range of people from developing country governments, bilateral aid agencies, multilateral institutions, foundations, and research centers, we came to the conclusion that $200 per assessed completer would be appropriate. This amount is large enough so that the expected funds would be significant relative to the government's education budget and existing foreign aid flows. It would get the attention of policymakers and engage them in improving efficiency. And it would provide enough resources to expand existing programs (school construction, teacher training, additional school bus routes) or create new ones (school meals, paved rural roads).

Although the payment could be negotiated in several ways, a very promising approach would be for multiple interested funders and recipients to agree on a single global price, perhaps in the context of a big push to accelerate progress toward the education Millennium Development Goal. This would uphold the notion of equality among beneficiaries and simultaneously provide lower income countries with more resources since their domestic costs are lower (see box 2.1).

A practical consideration: initial payments for implementing testing

For most countries, implementing a new test will be financially challenging. For practical reasons, then, it makes sense for funders to compensate the recipient for the cost of testing in the first few years. Based on international learning assessments, these costs could range between $5 and $25 per student.[9] We recommend that the COD Aid agreement in a typical low-income country include paying $20 for testing of each assessed completer up to the number of students enrolled in the last year of primary school when the agreement is signed. This payment would be phased out over five years. The number of assessed completers in excess of the base year enrollment would, of course, be compensated at the higher rate of $200 per assessed completer (box 3.3).

Total sums must be substantial enough to provide an incentive

The sums provided for progress must be substantial enough to provide an incentive for the recipient country. COD Aid payments can be put in the context of overall education budgets and foreign aid with an example from a low-income developing country in Sub-Saharan Africa. It has a population of about 35 million and a net enrollment in the last year of primary school of 64 percent of the children in the same age cohort. A payment of $200 per additional assessed completer in this case would imply annual transfers of between $10 million and $30 million if the recipient country reached 100 percent completion (box 3.4 and table 3.1). After this, the payments would gradually decline (see box 3.3 for details on the payment formula). This sum would still represent a fairly small share of total education spending in the country and only about 10–20

BOX 3.3

Formula for calculating the payment

The key elements of the payment formula in the text are:

1. Payments of $200 are made for each additional assessed completer.
2. The number of additional assessed completers is calculated as the difference between the actual number of assessed completers and the baseline.
3. The baseline is equal to base year enrollment (the number of assessed completers in the program's first year) during the program's first five years. The baseline is adjusted thereafter to be the number of assessed completers five years earlier.
4. Payments to implement testing are made during the first five years at the rate of $20 for each assessed completer and are phased out over this period.

To express this formally, during the first five years, the formula for payment for the additional assessed completers in year t would be:

$$\$200 * (AC_t - AC_b), \text{ when } AC_t > AC_b$$

where AC_t is the number of assessed completers in year t and AC_b is the number of assessed completers in the first year of the contract.

Assuming that, after five years, the country is testing more students than five years earlier ($AC_t > AC_{t-5}$), the formula for payments would then change to:

$$\$200 * (AC_t - AC_{t-5})$$

The payment for implementing testing can be expressed formally as:

$$P * AC_t$$

where AC_t is the number of assessed completers in year t, as before, and P is a payment that declines over five years: $20 per assessed completer for years 1 and 2, $15 for year 3, $10 for year 4, and $5 for year 5.

An illustration of the amounts of money involved for a typical developing country is shown in figure 1 in box 3.4 and in table 3.1.

percent of foreign aid to education (in African countries with per capita incomes below $5,000, primary education spending ranges from $20 to $800 per student).

In our view, the small amounts are likely to provide substantial incentives for at least two reasons. First, they would be attractive because they can be used flexibly, unlike the current largely committed spending for teacher salaries and other costs. Second, the public nature and transparency of the COD Aid contract means that all funds create reputational risks (and opportunities) for political leaders. In any event, the suggested size of the payment is meant to provide a starting point for funder-recipient negotiations; the actual payment specified in the contract would be the outcome of those negotiations.

Reporting and verification

Transparent reporting and verification are essential to the COD Aid agreement

Reporting and verification through audits ensure mutual accountability, provide incentives to improve education data management systems and enable social audits (box 3.5).[10] Constituents of both funders and recipients will find COD Aid agreements

BOX 3.4

Illustration of COD Aid payments to a developing country

To illustrate the financial implications of a hypothetical COD Aid contract, we applied the payment formula (see box 3.3) to data characteristic of a low-income Sub-Saharan country with a population of about 35 million. About 330,000 graduate each year from primary school, or almost two-thirds of the 517,000 children in the same age cohort.

The assumption that completion will increase at a rate between 2 and 3 percent higher than the historical average yields an annual COD Aid payment of $4 million in the agreement's second year (predominantly covering the cost of testing for the baseline enrollment), rising to $30 million at its peak after 13 years and declining to zero after 22 years, when all students in the slowly increasing age cohort are completing primary school (see figures 1 and 2). For comparison, funders provided the country with about $30 million for basic education and $70 million for the entire education sector in 2005.

Source: Table 3.1.

FIGURE 1

Progression of COD Aid payments over 22 years

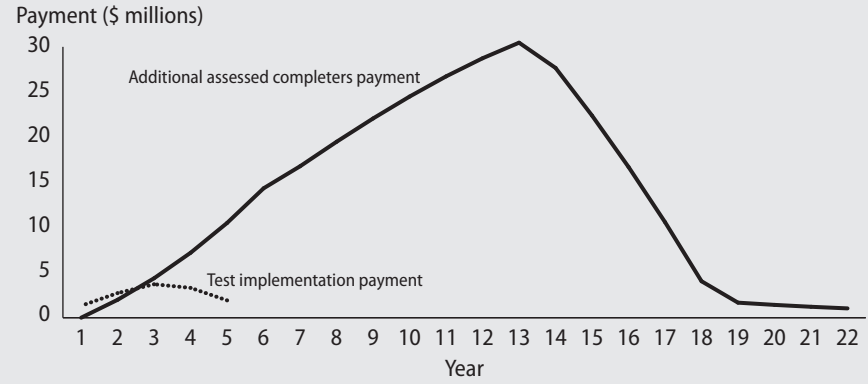


FIGURE 2

Number of enrolled students in final year of primary school

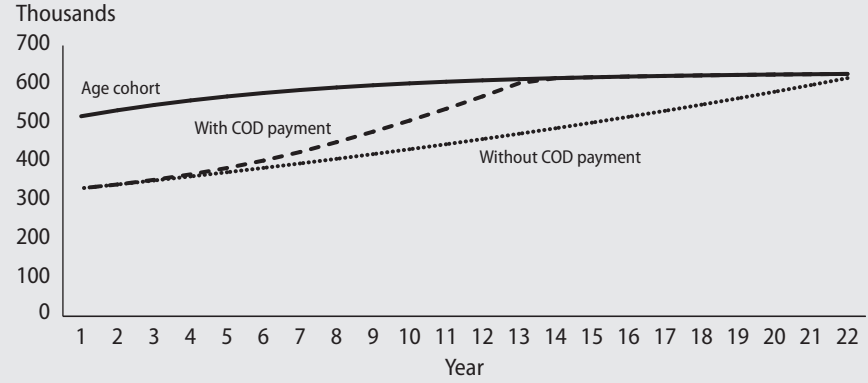


TABLE 3.1

Illustration of COD Aid payments to a developing country (thousands)

Year	Age cohort	Predicted final year enrollment		Baseline minus five years	Assessed completers[a]	Additional assessed completers[b]	Additional assessed completers payment ($)	Test implementation payment ($)	Total payment ($)
		Without COD Aid	With COD Aid						
0		327.0							
1	517.2	331.0	331.0	331.0	66.2			1,324	1,324
2	532.7	340.9	340.9	331.0	136.4	9.9	1,986	2,727	4,713
3	546.3	351.2	352.9	331.0	247.0	21.9	4,373	3,705	8,078
4	558.1	361.7	367.0	331.0	330.3	36.0	7,195	3,303	10,498
5	568.4	372.5	383.5	331.0	383.5	52.5	10,498	1,917	12,416
6	577.3	383.7	402.7	331.0	402.7	71.7	14,333		14,333
7	585.0	395.2	424.8	340.9	424.8	83.9	16,776		16,776
8	591.6	407.1	450.3	352.9	450.3	97.4	19,488		19,488
9	597.3	419.3	477.3	367.0	477.3	110.3	22,068		22,068
10	602.2	431.9	506.0	383.5	506.0	122.5	24,493		24,493
11	606.4	444.8	536.3	402.7	536.3	133.7	26,730		26,730
12	610.0	458.2	568.5	424.8	568.5	143.7	28,736		28,736
13	613.0	471.9	602.6	450.3	602.6	152.3	30,461		30,461
14	615.6	486.1	615.6	477.3	615.6	138.3	27,666		27,666
15	617.9	500.7	617.9	506.0	617.9	111.9	22,384		22,384
16	619.8	515.7	619.8	536.3	619.8	83.5	16,694		16,694
17	621.4	531.2	621.4	568.5	621.4	52.9	10,583		10,583
18	622.8	547.1	622.8	602.6	622.8	20.2	4,038		4,038
19	624.0	563.5	624.0	615.6	624.0	8.3	1,665		1,665
20	625.0	580.4	625.0	617.9	625.0	7.1	1,419		1,419
21	625.8	597.8	625.8	619.8	625.8	6.0	1,209		1,209
22	626.6	615.8	626.6	621.4	626.6	5.2	1,030		1,030

a. Assumes that the test will be phased in gradually over five years, testing 20 percent, 50 percent, 70 percent, 90 percent, and 100 percent of the enrolled students each successive year and 100 percent thereafter. For the first five years, the baseline equals the enrollment from year one; from year six on, the baseline equals the enrollment from five years prior.

b. The number of assessed completers minus the baseline. Final year enrollment is used for years one through four, before testing covers all students.

Source: Authors' calculations using data from a Sub-Saharan African country with 64 percent primary completion rate in 2005.

BOX 3.5

What role can social audits play?

COD Aid emphasizes that funds should be disbursed against a specific measurable outcome. The main risk is that the incentive would encourage the recipient to make decisions that achieve the measurable outcome at the expense of other important goals. In education, the primary concern is that basing payments on the number of students could privilege quantity over quality.

Rather than weakening the incentive of a single clearly measurable outcome, our judgment is that additional instruments should be found to mitigate this risk. The most promising avenue is to rely on and strengthen a country's mechanisms for ensuring public accountability. For example, we propose that a COD Aid program for education pay for the number of students who take a test and simultaneously require that test scores be publicly disseminated. This is a necessary, though not sufficient, condition for civil society to pressure government to maintain and improve the quality of education.

This mechanism is often referred to as a social audit—information about public services is periodically collected, processed, and disseminated to strengthen public accountability and improve services. This is an ongoing process in any democratic country (publishing government budgets, reporting to legislatures on public services). But initiatives to explicitly introduce social audits in developing countries have met with varying success.

- In Bangalore, India, the Public Affairs Centre developed a citizen report card to measure citizen satisfaction with public services. Publicly disclosing and debating the citizen report cards led several public agencies to improve their services.[1]
- Community-based monitoring of health care services in Uganda had a demonstrable effect on improving health outcomes.[2]
- A Mexican nongovernmental organization established to review public budget information discovered that 30 million pesos earmarked for women's health programs had been diverted to purposes contrary to the government's HIV/AIDS prevention policies. The funds were later returned to their intended use.[3]
- A South African nongovernmental organization used budget data to document disparities and declines in child support grants. Subsequent public pressure led the government to increase and reapportion funding.[4]

Social audits are not foolproof, and other experiences demonstrate conditions under which they may fail to generate government responses (Banerjee, Banerji, and others 2008). But for trying a new aid modality such as COD Aid, experience shows that including mechanisms to facilitate social audits is a promising avenue.

Notes

1. Ravindra 2004.
2. Björkman and Svensson 2007.
3. Ramkumar 2008 [add to references].
4. de Renzio and Krafchik 2007.

much more credible if they have access to easily understood and reliable information. COD Aid contracts should therefore clearly specify the reporting and verification that must be fulfilled before payment is made.

Public reports of progress will then be used by the funder for calculating payments, and by a broader audience of policymakers, legislators, media, civil society, and parent groups. Through reports on test taking and test scores, they will begin to

understand, interpret, and engage public policymakers, holding them accountable for the education system's performance in improving access, quality, and equity. Public disclosure of the verification reports would also allow civil society organizations and the public to assess the government's integrity and to pressure schools or administrators who fail to report or who manipulate information.

The reporting and verification also provide incentives to improve education data. In most developing countries, education system information is unreliable. For example, despite considerable investment in South Africa's Education Management Information System, the country still has no agreement on the actual number of schools in the country, let alone the number of students completing primary school. In such circumstances, the fact that the COD Aid agreement pays against officially reported information subject to an audit is likely to stimulate better recordkeeping.

The COD Aid contract requires four levels of disclosure.
The recipient would be expected to publicly disclose information at four levels:

- *Information on the test design and administration according to current professional standards.* This would allow technical experts to debate and judge the validity and reliability of the test and to offer comments to improve the national learning assessment.

- *The number of assessed completers and average test scores at a level of disaggregation specified in the contract (state, municipal, district, or school).* The disaggregation should be to the smallest level possible, given the test's validity, reliability, and precision. Because the national learning assessment will probably improve over time, the funder and recipient might agree to a fairly broad aggregation that becomes more refined.

- *The raw data to researchers, taking appropriate measures to protect the privacy of individuals who took the tests.* Providing open access to raw data increases the chances that worthwhile analysis will emerge, that suggestions for improvement will arise, and that errors will be detected. (We discuss country research on COD Aid in chapter 5.)

- *The results of the audit, as discussed below.*

Paying against officially reported information subject to an audit is likely to stimulate better recordkeeping

The role of an audit
The independent audit of the number of assessed completers should be the subject of a separate contract between funders and a third-party auditor (see the model contract in the appendix). The independence of the

audit, contracted by the funder, is a critical part of the agreement.[11] The audit provides all parties with greater confidence in the outcome measures and assists the country in identifying any problems with its information system so that improvements can be made. To detect problems, the audit requires that each student have a unique identifying number, that the auditor apply a retest in a random sample of schools, and that consequences for discovered discrepancies are clearly specified.

Unique identification is required for an accurate audit. The first requirement for an accurate audit is to track who is and is not eligible to take the test. At a minimum, this requires that each assessed completer be assigned a unique identification number. Preferably, the official testing system would use a national identification number to identify each student. Numbers for national identification cards would serve such a purpose without requiring a substantial investment. With a unique identifier, the audit can detect double counting, repeated test taking, and other problems that could lead to inflated reports. If a country introduced identification numbers for students as a part of the COD Aid agreement, this would also benefit data management. For example, in some developing countries, students whose families migrate from one part of the country to another may be registered as dropouts in their original schools even though they have enrolled elsewhere, leading to inaccuracies that make management of the education system more challenging. A national identification number system would help address this.

Retesting a random sample of schools. To verify the number of assessed completers, the audit would include not only cross-checks of student identification numbers, but also a retest in a random sample of schools within a fairly short period of time after the official test (such as one or two weeks).[12] The retest would require assembling a team that can visit a number of schools throughout the country and administer a retest for all eligible children in those schools. In addition, the team would check identification numbers against electronic databases and inspect enrollment records. This retest could detect whether tests were being applied properly and reduce cheating by comparing test scores. It could also identify ineligible test takers and problems with enrollment records.

The audit provides all parties with greater confidence in the outcome measures

To maximize attendance for the retest, the date of the retest would be publicized, but no one would know beforehand which schools were to be retested. Provisions for test security would ensure that test records are not tampered with after the retest.

Although the actual cost and scale of the audit would vary across countries, data from current international tests suggest that estimates with sufficient precision can be derived from retesting a sample of about 180 schools, involving about 9,000 students (box 3.6). The official reports would be accepted as accurate if the reported test scores

BOX 3.6

Cost and resources required for the audit

The sample size necessary to estimate average test scores for each school with a confidence interval of 5 percent depends on average test scores and their variance. Experience with international tests suggests these numbers might be 250 for average test scores with a standard deviation of 100. Creating a one-sided 95 percent confidence interval for the average school results with a width of 12.5 (5 percent of 250) would then require a sample size of about 180 schools. Assuming an average of 50 students per school would imply retesting about 9,000 students, which is close to the size of sample-based testing used by the Programme for International Student Assessment.

Field work would require about two person-days per school to administer the test and inspect school or district enrollment records. Further analysis would be required to derive useful results and report them in an understandable form. The cost of such an audit would be about $250,000 in addition to the basic cost of developing and applying the official test.

Source: Crouch and Mitchell 2008.

are less than a threshold calculated as a one-sided confidence interval derived from the retest sample.

The total cost could be as low as $250,000 if the audit can rely on an existing system of unique identification numbers (such as a national identity card system), reasonable data on how many primary schools the country has and where they are located, and access to an existing test that meets the necessary standards. Should it be necessary to create an identification number, conduct a census of primary schools, or develop a test without any prior groundwork, the costs would be correspondingly higher. Even with these additional costs, however, the total expenditure would be modest relative to the amount of aid and to the benefits that the audit provides. The audit is also the only part of the COD Aid agreement that entails administrative costs, unlike input-based aid, which requires administrative inputs throughout the entire process (see figure 2.1).

In sum, the main requirements for the audit would be to invest in an identification number allocation system (or adopt an existing one), set up an electronic student database, develop tests, and manage the logistics for widespread and secure test administration, scoring, and reporting. The main effort of the audit would involve retesting a sample of schools, inspecting records, analyzing data, and generating easily understood reports.

Consequences: do the official reports pass or fail? Funders and recipients will expect the audit to confirm that the official reports are accurate. But it is very important for the agreement to specify precisely what constitutes accuracy and what happens if the audit detects problems. It should be kept in mind that the audit will be estimating only the total number of

The total cost of the audit could be as low as $250,000

The results of the audit will also be public information

assessed completers and their average test scores within a statistical margin of error. From our consultations with experts on this matter, we propose that the official report be considered accurate if the reported number of assessed completers and the average test scores are no more than 5 percent higher than the auditor's upper-bound estimate using a 95 percent level of statistical significance.

Once the official report passes the audit, the funder will make the payments specified in the contract. In our proposal, this means paying $20 per assessed completer up to the number of students who were enrolled in the last year of primary school in the first year of the COD Aid agreement (with the payment being phased out gradually over subsequent years) and an additional $200 per assessed completer for all students beyond the initial enrollment.

Consequences have to be specified for the possibility that the official reports overstate either the number of assessed completers or the average test score. In our proposal, the consequences are simple and direct. If the number of assessed completers is overstated, the funder will still make payments to the recipient but only according to the auditor's point estimate of assessed completers. If the test scores are overstated, the payment per assessed completer will be reduced by a factor proportional to overstatement. Structuring the penalties in this way lowers the stakes. The recipient is very likely to receive some payment, rather than all or nothing. But the penalty design gives the recipient an incentive to report the information as accurately as possible (see the appendix for further discussion of auditing and consequences).

Consequences: public audit reports create pressure for accurate reporting. The results of the audit will also be public information. The information must be made available in forms that are useful to the public and at the greatest level of disaggregation that is appropriate and feasible. The funder-recipient and funder-auditor contracts should specify what data will be public. For example, audit reports could list the names of schools whose retest numbers and scores diverged significantly from the official report. This aspect of the audit could encourage local accountability and stimulate attention to other aspects of schooling that cannot be captured in one or a few index numbers. The audit could also provide parents and civil society organizations with information about equity, quality of instruction, data manipulation, or unintended consequences, equipping them to pressure the education system to improve.

Provision for contingencies

COD Aid agreements are contracts that establish a payment for delivery of a particular good or service. As with any contract, it is impossible to foresee the future and anticipate all developments that might affect the parties' ability to fulfill their

obligations. Such concerns can be addressed by designing an agreement that is robust in the face of predictable areas of disagreement, with contingencies for foreseeable events and arbitration for unforeseeable developments.

The proposed design for COD Aid to primary education incorporates features to make it robust. Choosing an indicator that is relatively easy to verify limits opportunities for manipulation and disagreement over payments. Measures to ensure transparency and public oversight also strengthen the agreement by reducing the likelihood that either party will try to renegotiate without good justification.

Contingencies are still necessary, however, for events beyond the recipient's control that could interfere with schooling or disrupt testing—such as major natural disasters or declines in the world price of an important export. While countries facing such crises may require additional aid, it is important for those funds to come through other channels. We think that appropriate contingencies could be included in the contract to allow rescheduling tests or delaying reports, but that payments under the COD Aid agreement should not otherwise be adjusted. Standard conditions for either party to withdraw from the contract are also necessary, as is a process for arbitration if irresolvable differences emerge. Recourse to arbitration should involve costs, so that it is not undertaken frivolously.

COD Aid is feasible

Our view, in short, is that the key features of a COD Aid agreement—payment for outcomes, hands-off funders and responsible recipients, independent verification with transparency and complementarity—can be captured in a robust contract reflecting a variety of specific country settings. We have shown how a particular indicator of progress—the number of additional assessed completers—can serve as the basis for a COD Aid agreement and how levels of payment, verification procedures, and contingencies could work in practice. To explore these ideas in greater detail and provide assistance for developing a COD Aid program, we have included term sheets for drafting the necessary contracts in the appendix. Chapter 4 considers issues related to funding and implementing the agreement once it has been negotiated and signed.

Conclusion: keeping it simple

This chapter has covered a great deal of detail, from base year enrollment and additional assessed completers to payment formulas, testing methods, and auditing requirements. The next chapter will discuss funding and implementation at a similar level of detail. This precision and specificity is essential to making a COD Aid agreement work. But specific need not mean complex. Indeed, in our review of lessons from previous efforts to reform foreign aid and our extensive consultations, we repeatedly returned

The key features of a COD Aid agreement can be captured in a robust contract

to the imperative of keeping it simple. We conclude this chapter by underscoring four principles of simplicity:

- Choose a simple indicator.
- Choose a simple incentive.
- Support other ways of verifying progress.
- Use existing expertise.

Choose a simple indicator

In developing a COD Aid agreement, it is very easy for the discussion to begin with a simple indicator—which by definition can only approximate the true goal of a program—and move rapidly toward sophisticated indicators. Using sophisticated indicators can, however, undermine qualities of a COD Aid agreement that are critical to its success. For an agreement to be credible, the indicator must be clearly defined, focused, measurable with sufficient precision, and verifiable to reduce uncertainty over the deliverable. Credibility, in turn, is necessary for the agreement to create an incentive for the recipient. And unless the indicator can be easily explained to the public, it will be more difficult to hold funders and recipients accountable for compliance. This does not mean that a simplistic indicator should be chosen—it means that the tendency toward sophistication should be tempered by practicality and ease of understanding.

Choose a simple incentive

Another temptation in developing a COD Aid agreement is to take the concept of an incentive to extremes. The basic incentive is a payment for improvements in a measurable outcome. But if a single incentive is good, wouldn't many incentives be better? The temptation arises to add further and more complex incentives, such as rewarding complementary goals (improving teacher incentives, increasing the number of schools with a certain number of textbooks, introducing school autonomy, strengthening management information systems). It is also tempting to elaborate on the payment structure to closely reflect the amount of effort and expenditure required at different stages (such as successively increasing the per-student payment to address the increasing difficulty of reaching more marginalized children). As with choosing an indicator with the right balance of simplicity and sophistication, the incentive design must also find an appropriate balance. An incentive that involves different levels of payment for many related indicators is less effective than a single payment for progress toward a single outcome indicator because the added complexity necessarily diffuses managerial and political attention and reduces transparency to the public and civil society.

Choosing an indicator with the right balance of simplicity and sophistication

Support other ways of verifying progress

While the COD Aid agreement itself needs to be simple, this does not preclude encouraging other agencies, civil society groups, or research institutions to monitor the full range of desirable outcomes. Our example focuses on a specific measurable indicator: the number of students completing school and taking a test. But even that indicator is only a proxy for the true objective: a well rounded education.

While the COD Aid agreement needs to be simple, other agencies, civil society groups, or research institutions could monitor the full range of desirable outcomes

In countries with effective civil society organizations, groups can be encouraged to monitor inputs to learning and other aspects of schooling, such as teacher absenteeism, student attendance, and nuanced quality measures of classroom instruction and student achievement. Equipped with such information about a fuller range of indicators, civil society groups can exert pressure on the government to maintain and improve quality, and professional associations can engage in informed discussions with education policymakers. By encouraging the measurement and dissemination of other aspects of schooling, the COD Aid indicator can remain simple while civil society groups and other institutions provide more sophisticated oversight on aspects less easily reduced to single index numbers.

Use existing expertise

Funders and recipients have been working in most development sectors for decades, and a large body of knowledge and expertise has been generated through those experiences. These experts are an indispensable resource in solving the problems in designing a practical and useful COD Aid agreement. Indeed, the proposal described here is the outcome of many problem-solving sessions with international education experts over two years—and of background papers and notes commissioned to address particularly important issues such as testing, data systems, and auditing.[13]

Notes

1. Levine, Birdsall, and Ibrahim 2003.
2. This proposal comes from a background paper by Crouch and Mitchell (2008) commissioned for this project. Further arguments in favor of this indicator can be found in that background paper.
3. This is the definition used by the World Bank in its EdStats for the numerator in its primary completion rate. The United Nations Educational, Scientific, and Cultural Organization reports the same statistic but more transparently calls it gross intake ratio in the last grade of primary. See also Bruns, Mingat, and Rakotomalala (2003), p. 40.
4. This section draws heavily on a background paper by Lockheed (2008) commissioned for this project.
5. Lockheed 2008; Holland and Rubin 1982; Linn 2005.

6. Available international and regional tests include the Organisation for Economic Co-operation and Development's Programme for International Student Assessment or the Southern African Consortium for Monitoring Educational Quality.
7. Such organizations include Trends in International Mathematics and Science Study, Progress in International Reading Literacy Study, and Programme d'Analyse des Systèmes Educatifs de la CONFEMEN.
8. Drawn from Crouch and Mitchell (2008).
9. Lockheed 2008.
10. This section draws directly from and extensively quotes a background paper by Crouch and Mitchell (2008) commissioned for this project.
11. See box 7.1 for an example of how relying on the recipient to report progress can create problems.
12. The sample must be truly random, such as including schools that are difficult to reach because of distance, conflicts, or transportation difficulties. Otherwise, there would be a temptation to manipulate results in schools unlikely to be visited.
13. A list of these background papers and notes is available on CGD's website at www.cgdev.org/section/initiatives/_active/codaid/papers_and_resources.